面向视觉跟踪的目标特征表达研究

韩守东[1,2,*]　陈永志[1]　黄飘[1]　陈阳[1]
[1]（华中科技大学自动化学院多谱信息处理技术国家级重点实验室 武汉 430074）
[2]（深圳华中科技大学研究院　深圳　518057）

摘要： 视觉跟踪是计算机视觉中的一个重要课题，近年来随着它在无人机、智能交通、人机交互等方面的深入应用，受到了研究者的广泛关注与研究，并取得众多的研究成果。在视觉跟踪中，目标特征表达作为最基础也是最重要的部分，直接关系着最终的跟踪效果，合适的目标特征表达可大大提升结果的准确性。本文对视觉跟踪中常用的目标特征表达方法进行了整理与归纳，将其分为了三类，分别为视觉特征、数学特征和以卷积神经网络为主的语义特征，并对它们进行了详细地分析和比较。其中，相对于视觉特征和数学特征，语义特征能够更加有效地体现物体内在类别信息，对形变、遮挡等均有较高的鲁棒性。同时，本文还提出了目标特征表达在复杂性、准确性和鲁棒性方面所存在的问题。

关键词： 视觉跟踪；目标特征；特征表达；视觉特征；数学特征；语义特征
分类号：TP391.4

Review on target feature expression for visual tracking

Han Shoudong[1,2,*]　Chen Yongzhi[1]　Huang Piao[1]　Chen Yang[1]
[1](National Key Laboratory of Science and Technology on Multispectral Information Processing, School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China)
[2](Research Institute of Huazhong University of Science and Technology in Shenzhen, Shenzhen 518057, China)

Abstract: Visual tracking is an important subject in the field of computer vision, recently with the in-depth application in unmanned aerial vehicle, intelligent transportation, and human-computer interaction, it has been attracting researchers' increasing attentions and researches, and great research achievements have been made. As the most fundamental and important part of visual tracking, the target feature expression is directly related to the tracking effect. The appropriate target feature expression can greatly improve the accuracy of tracking results. In this paper, the target feature expression methods commonly used in visual tracking are collated and summarized, and here they are classified into three categories, such as visual feature, mathematical feature, and semantic feature mainly based on convolutional neural network. Additionally, these three categories are analyzed and compared in detail, which concludes that relative to the visual feature and mathematical feature, the semantic feature can more effectively reflect the intrinsic classification information of target, and has high robustness for target tracking with shape change and partial occlusion. Moreover, this paper proposes some problems of target feature expression existing in complexity, accuracy, and robustness.

收稿日期：2018年-1月-15日；**收到修改稿日期：**年-月-日
基金项目：国家自然科学基金（61105006）；国家863计划（2015AA015904）；深圳战略新兴产业发展专项基金（JCYJ20170307172130906）
作者简介：韩守东（1983-），男，博士，副教授，硕士生导师，主要从事视觉目标检测与跟踪、图像/视频分割与合成、深度学习算法与应用、人脸识别与3D建模等方面的研究。E-mail：shoudonghan@hust.edu.cn
*通信联系人。E-mail:shoudonghan@hust.edu.cn

Key words: visual tracking; target feature; feature expression; visual feature; mathematical feature; semantic feature

1 引 言

视觉跟踪是计算机视觉的热点研究课题，在无人机 Xiao L, Meng G, Luo H, et al. Dynamic path planning based on improved boundary value problem for unmanned aerial vehicle[J]. Cluster Computing, 2016, 19(4): 1-10.Wang X, Li B, Geng Q. Runway Detection and Tracking for Unmanned Aerial Vehicle Based on an Improved Canny Edge Detection Algorithm[C]. International Conference on Intelligent Human-Machine Systems and Cybernetics, 2012, 149-152.、智能交通 Hadi R A, George L E, Mohammed M J. A Computationally Economic Novel Approach for Real-Time Moving Multi-vehicle Detection and Tracking toward Efficient Traffic Surveillance[J]. Arabian Journal for Science & Engineering, 2016, 1-15.Hai D, Hua T. Development of a tracking-based system for automated traffic data collection for roundabouts[J]. Journal of Modern Transportation, 2017, 1-12.、人机交互 Joslin C A E Q. Dynamic gesture recognition[C]. Instrumentation and Measurement Technology Conference, 2005. 1706-1711.等方面获得了成功的应用，受到研究者的广泛关注与深入研究。对于视觉跟踪的步骤，Naiyan WangWang N, Shi J, Yeung D, et al. Understanding and diagnosing visual tracking systems[C]. The IEEE International Conference on Computer Vision (ICCV), 2015. 3101-3109.将跟踪过程分为五个部分：运动模型、特征提取、观测模型、模型更新和总体效果后处理；Yi WuWu Y, Lim J, Yang M. Online object tracking: A benchmark[C]. IEEE Conference on Computer Vision and Pattern Recognition((CVPR)), 2013. 2411-2418.认为跟踪主要包括目标表示、搜索方法和模板更新这三个部分；Xi LiLi X, Hu W M, Shen C H, et al. A survey of appearance models in visual object tracking[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2013, 4(4): 51-58.则将跟踪过程分为目标初始化、特征建模、运动估计和目标定位这四个步骤。虽然不同研究者对于跟踪步骤有不同看法，但是其共性都包含了目标特征表达（特征提取、目标表示、特征建模）这部分，可见目标特征表达在视觉跟踪中的重要性。目标特征表达直接关系着跟踪效果，选择合适的目标特征表达方法，可大大提升结果的准确性。

理想的目标特征表达应具有可重复性、可区分性以及高效等特性，能够对图像亮度变化、尺度变化、旋转和仿射变换等足够鲁棒。本文将目标特征表达方法分为三种：视觉特征、数学特征和语义特征。视觉特征是一种类似于人眼看到的直观特征表达，数学特征是将图像信息看成一种抽象的数学信号来构建目标的内在数学联系，而语义特征则是表达物体内在的类别信息。表 1 列出了本文将要描述的所有目标特征表达方法。

表 1 目标特征表达分类

Table 1 Classification of Target Feature Expressions

特征分类	特征含义	具体特征举例	详细分类
视觉特征	视觉的直观特征表达	颜色特征、光流特征、HOG 特征、SIFT 特征、	第 2 节

		纹理特征等	
数学特征	构建目标内在数学联系	高斯混合模型、子空间模型、WSL 混合模型等	第 3 节
语义特征	表达物体内在类别信息	卷积神经网络等	第 4 节

为了对目标特征表达方法进行全面地认识，本文对视觉跟踪中常见的目标特征表达方法进行了归纳与整理。本文的结构安排如下：第 1 节是引言，第 2~4 节分别详细地介绍了视觉特征、数学特征和语义特征，并对其进行了详细地分析和比较。第 5 节对本文进行了总结，并提出了目标特征表达方法在复杂性、准确性和鲁棒性方面所存在问题。

2 视觉特征

视觉特征是从图像中提取出来的用以表征目标的某种视觉信号。根据视觉特征表征目标的范围，可分为全局视觉特征和局部视觉特征。全局视觉特征反映的是目标表观的全局信息，主要包括颜色特征、光流特征等。而局部视觉特征主要提取感兴趣点或者用显著性检测的方法来编码目标的表观信息，主要包括 HOG 特征、SIFT 特征、纹理特征等。在一些复杂场景中，一种视觉特征可能难以有效地表征目标，因此有时需要将多种视觉特征进行融合。下面本文将从全局视觉特征、局部视觉特征、融合视觉特征这三个方面对常用的视觉特征进行简述。

2.1 全局视觉特征表达

（1）颜色

颜色特征是各类视觉跟踪算法中应用最为广泛的视觉特征，它描述了图像中所包含物体或场景的表面属性。在计算机视觉中，描述视觉目标的颜色空间有很多种，如 RGB（Red 红、Green 绿、Blue 蓝）颜色空间黄国祥. RGB 颜色空间及其应用研究[D]. 中南大学, 2002: 88.、HSI（Hue 色调、Saturation 饱和度、Intensity 亮度）颜色空间庞晓敏, 闵子建, 阚江明. 基于 HSI 和 LAB 颜色空间的彩色图像分割[J]. 广西大学学报(自然科学版), 2011, 36(6): 976-980.、HSV（Hue 色调、Saturation 饱和度、Value 明度）颜色空间 Bradski G. Real time face and object tracking as a component of a perceptual user interface[C]. Applications of Computer Vision, 1998, 214-219.、Lab（Luminosity 亮度）颜色空间 Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]// Computer Vision and Pattern Recognition. IEEE, 2010:2544-2550.等，不同的颜色空间其稳定性不同。不过，J.Ross Beveridge 等人在相关滤波类目标跟踪算法中直接使用了灰度特征 Danelljan M, Khan F S, Felsberg M, et al. Adaptive Color Attributes for Real-Time Visual Tracking[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014:1090-1097.，虽然保留了所有颜色特征，但是也包含了很多噪声。另外值得注意的是，Martin Danelljan 首次在目标跟踪领域引入 Color Attributes(black、blue、brown、grey、green、orange、pink、purple、red、white、yellow)共 11 维颜色特征，并将其改进为 10 维 CN 特征，与此同时，其在跟踪过程中利用 PCA 技术将 CN 特征降维至 2 维 Choi J, Chang H J, Yun S, et al. Attentional Correlation Filter

Network for Adaptive Visual Tracking[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017.。并且作者也在文中证明了CN颜色特征在目标跟踪领域中的优势。

由于单纯的像素级颜色特征易受噪声干扰影响 Kengyew S, Kittler J, Petrou M. Defect detection in random colour textures[J]. Image and Vision Computing, 1996, 14(9): 667-683.，颜色直方图包含颜色统计信息，能降低噪声的影响，在跟踪中最为常见贾慧星, 章毓晋. 基于梯度方向直方图特征的多核跟踪[J]. 自动化学报, 2009, 10): 1283-1289.。Bradski等人 Bradski G. Real time face and object tracking as a component of a perceptual user interface[C]. Applications of Computer Vision, 1998, 214-219.在HSV颜色空间中使用颜色直方图来表示目标，并在Camshift（Continuously Adaptive Mean Shift）框架中实现视觉跟踪。王晓卫等人王晓卫, 王旭东, 贺明. 基于直方图比的背景加权的 Mean Shift 目标跟踪算法[J]. 强激光与粒子束, 2016, 28(05): 19-23.在均值漂移（Mean Shift）Fukunaga K, Hostetler L D. The estimation of the gradient of a density function, with applications in pattern recognition[J]. IEEE Transactions on Information Theory, 1975, 21(1): 32-40.跟踪框架中使用目标和背景的颜色直方图比值进行视觉跟踪。颜色直方图的优点是可以有效获取目标区域中视觉特征的分布特性，并且不受图像旋转和平移变化的影响。但是由于颜色直方图只考虑了目标颜色的统计信息，丢失了目标颜色的空间信息，使得在跟踪过程中容易造成跟踪失误。因此有研究者提出了空间颜色直方图 Birchfield S T, Rangarajan S. Spatiograms versus histograms for region-based tracking[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR), 2005. 1158-1163.，通过在直方图的计算过程中引入空间位置信息以获得较好的跟踪效果。一般可采取的策略有两种，一种是联合“空间-颜色”建模，采用（x，y，R，G，B）来描述位置和颜色的联合目标状态，其中（x，y）为位置信息，（R，G，B）为颜色信息。另一种策略则是分块策略，即先将目标区域分割成若干块，然后将目标区域的空间信息融合到目标特征表达中。

虽然颜色视觉特征对于处理部分遮挡、尺度变化和旋转平移等视觉跟踪问题具有足够的鲁棒性，但在复杂场景中，如光线强度变化、目标姿态变化或者场景中出现相同颜色干扰时，都会导致误跟踪。因此在复杂场景中，单一地采用颜色特征进行视觉跟踪，跟踪效果可能不太理想。

（2）光流

在图像序列中只要有运动的存在，就会有光流特征产生。光流反应了图像上每一像素点的灰度变化趋势，不仅包含图像中目标的运动信息，而且包含了三维物理结构信息。光流特征在视频分割和视觉跟踪中应用广泛裴巧娜. 基于光流法的运动目标检测与跟踪技术[D]. 北方工业大学, 2009: 69.王亮. 光流技术及其在运动目标检测和跟踪中的应用研究[D]. 国防科学技术大学, 2007: 78.，一般在跟踪领域中常用的光流法有连续光流法 Nagel H, Enkelmann W. An investigation of smoothness constraints for the estimation of displacement

vector fields from image sequences[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(5): 565-593.Kearney J K, Thompson W B, Boley D L. Optical flow estimation: An error analysis of gradient-based methods with local optimization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(2): 229-244.和特征光流法 Weng J, Ahuja N, Huang T S. Matching two perspective views[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(8): 806-825.。

光流特征的核心是计算运动目标的速度，在计算过程中存在灰度不变性原理假设 Delpiano J, Jara J, Scheer J, et al. Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy[J]. Machine Vision and Applications, 2012, 23(4SI): 675-689.：假定视频图像中固定点的灰度值瞬时不变。光流特征计算常用的方法有微分法和块匹配法 Delpiano J, Jara J, Scheer J, et al. Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy[J]. Machine Vision and Applications, 2012, 23(4SI): 675-689.，由于微分法必须要求图像可微，且当在相邻图像帧之间的偏移量大时会产生较大误差，而块匹配法的运算量比较大，区域块大小的选择会对光流结果造成很大的影响。针对这些不足，一些学者提出了改进的光流法，比如立体匹配光流法胡庭波, 吴涛, 贺汉根. 基于立体匹配技术的光流场计算方法[J]. 计算机工程与科学, 2006, 28(10): 50-53.、特征光流法 Smith S M. ASSET-2: real-time motion segmentation and shape tracking[C]. Computer Vision, 1995. Proceedings, Fifth International Conference on, 1995. 80-85.Shi J, Tomasi C. Good features to track[C]. IEEE Computer Vision and Pattern Recognition(CVPR), 1994. 593-600.等。在视觉跟踪中，光流特征的主要问题在于大多数光流特征的计算耗时，不能满足实时跟踪的要求。因此很多跟踪方法都会通过结合光流特征和其它方法的方式来实现对目标的检测和跟踪，例如吴垠等吴垠, 李良福, 肖樟树, et al. 基于尺度不变特征的光流法目标跟踪技术研究[J]. 计算机工程与应用, 2013, 49(15): 157-161.提出先对图像求取 SIFT 特征点 Lowe D G. Object recognition from local scale-invariant features[C]. Seventh IEEE International Conference on Computer Visionthe, 1999. 1150-1157.，再通过计算特征点的光流特征来进行跟踪；陈添丁等陈添丁, 胡鉴, 吴涤. 稀疏光流快速计算的动态目标检测与跟踪[J]. 中国图象图形学报, 2013, 18(12): 1593-1600.采用 Harris 角点作为特征点，再结合光流特征来实现动目标检测与跟踪。

光流特征的优势在于可以获得目标的运动参数，不仅能够对多目标进行运动分析，还能在不知道跟踪场景具体信息的情况下，有效地检测出运动对象。但是利用光流特征进行视觉跟踪仍存在诸多局限性：在实时性方面，光流特征的计算非常耗时，很难实现实时目标跟踪；在具有复杂场景的实际应用中，光流计算的灰度不变性假设条件往往不能满足，从而不能求解出正确的光流特征；另外，光流特征容易受到噪声污染，对光照的鲁棒性较差，对于跟踪目标处于静止状态的情况不能处理。

（3）其他全局特征

边缘特征 Bowyer K, Kranenburg C. Edge detector evaluation using empirical ROC curves[C]. IEEE Computer Vision and Pattern Recognition(CVPR), 1999. 100-105.通过检测目标边界的强度变化来实现对目标的检测与跟踪。Canny 边缘检测算子 Wang X, Li B, Geng Q. Runway Detection and Tracking for Unmanned Aerial Vehicle Based on an Improved Canny Edge Detection Algorithm[C]. International Conference on Intelligent Human-Machine Systems and Cybernetics, 2012, 149-152.由于计算简单和精确度较高的特点，在视觉跟踪中应用非常广泛。协方差特征 Su Y Y, Zhao Q J, Zhao L J, et al. Abrupt motion tracking using a visual saliency embedded particle filter[J]. Pattern Recognition, 2014, 47(5): 1826-1834.能够收集多个矩阵样本之间的协方差，能够捕捉目标外观的关联信息，与大多数特征相比，协方差是一种便于计算的低维特征。主动轮廓特征（Snake）Sun X, Yao H, Zhang S. A novel supervised level set method for non-rigid object tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011. 3393-3400.采用带符号的距离映射来隐式地对目标的边缘信息进行编码，以水平集进化方法为基础，能够使用复杂的形状精确地分割出目标区域，可以用于非刚体目标的跟踪。另外，小波特征描述 Jepson A D, Fleet D J, El-Maraghi T F. Robust online appearance models for visual tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1296-1311.对位置误差、光线等因素具有强的鲁棒性，也是常用的全局特征。表 2 对上述各全局特征表达进行了归纳。

表 2 全局特征表达分类

Table 2 Classification of Global Feature Expressions

全局特征	参考文献	特点
颜色	[9-23]	对部分遮挡、尺度变化和旋转鲁棒；易受噪声的影响
光流	[24-35]	可获得目标运动参数，可以对多目标进行运动分析，以及运动对象检测；计算耗时，对光照鲁棒性差
边缘特征	[2][36]	能检测目标边界的强度变化，计算简单、精确度高
协方差特征	[37]	能够捕捉目标外观的关联信息
主动轮廓特征	[38]	能精确的分割出目标区域，可以用于非刚体目标的跟踪
小波特征	[39]	对位置误差、光线等因素具有鲁棒性

2.2 局部视觉特征表达

（1）HOG 特征

方向梯度直方图（Histogram of Oriented Gradient, HOG）是一种描述目标局部轮廓与形状的特征。与其他的特征描述方法相比，HOG 特征对图像几何和光学的形变都能保持很好的不变性，能够很好地表征图像局部像素点之间的关系，因此在各种检测和跟踪算法中被广泛应用。

HOG 特征最初应用在目标检测领域，DalalDalal N, Triggs B. Histograms of oriented gradients for human detection[C]. IEEE Computer Vision and Pattern Recognition(CVPR), 2005.

886-893.提出 HOG 特征与 SVM 分类器相结合的行人检测方法，取得了显著的检测效果。在视觉跟踪领域，HOG 特征如今已被各种算法广泛使用。Henriques 等 Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.提出的 KCF（Kernelized Correlation Filter）算法，Danelljan 等 Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking[C]. the British Machine Vision Conference, 2014. 524-533.提出的 DSST（Discriminative Scale Space Tracker）算法等，都使用了 HOG 特征，并取得了较高的准确率和较快的速度。

虽然 HOG 特征被广泛应用，但是 HOG 特征仍然存在计算较复杂、特征维度高、实时性较差、检测精度有待进一步提高等问题。因此，一些学者对 HOG 特征进行了选择与融合，以提高跟踪算法的准确性和实时性：文献田仙仙, 鲍泓, 徐成. 一种改进 HOG 特征的行人检测算法[J]. 计算机科学, 2014, 41(09): 320-324.使用 Fisher 挑选准则对 HOG 特征块进行筛选，保留下区分力较强的特征块进行行人检测，提高了算法的实时性。文献 Wu J, Yang S, Zhang L. Pedestrian detection based on improved HOG feature and robust adaptive boosting algorithm[C]. 2011 4th International Congress on Image and Signal Processing (CISP), 2011. 1535-1539.将 HOG 特征和 Haar 特征相结合，而文献 Sanin A, Sanderson C, Harandi M T, et al. K-tangent spaces on riemannian manifolds for improved pedestrian detection [C]. IEEE International Conference on Image Processing ICIP, 2012. 473-476.将局部保持投影（Locality Preserving Projection, LPP）特征和局部二值模式（Local Binary Pattern, LBP）Ojala T, Pietikainen M P, Maenpaa T M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.宋克臣, 颜云辉, 陈文辉, et al. 局部二值模式方法研究与展望[J]. 自动化学报, 2013, 39(6): 730-744.特征与 HOG 特征相结合，通过特征融合的方式进行视觉跟踪，取得了较好的跟踪效果。

（2）SIFT 特征

SIFT 特征即尺度不变特征变换（Scale-Invariant Feature Transform）特征，由 David Lowe 于 1999 年首次提出 Lowe D G. Object recognition from local scale-invariant features[C]. Seventh IEEE International Conference on Computer Visionthe, 1999. 1150-1157.，并在 2004 年完善而成 Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.。它是一种对目标的平移、尺度、旋转、亮度、仿射、光照等变化都较为鲁棒的描述子。

由 SIFT 算法提取出的 SIFT 特征点具有高度的特异性，能够保证算法匹配的精度，因此 SIFT 特征被广泛应用于视觉目标检测与跟踪中。Huiyu ZhouZhou H, Yuan Y, Shi C. Object tracking using SIFT features and mean shift[J]. Computer Vision and Image Understanding, 2009,

113(3): 345-352.将SIFT特征与Mean Shift相结合，并且用期望最大化算法（Expectation Maximization, EM）来评估目标的置信区域，获得目标位置；Saeid FazliFazli S, Pour H M, Bouzari H. Particle Filter based Object Tracking with Sift and Color Feature[C]. International Conference on Machine Vision (ICMV), 2009. 89-93.将SIFT特征与粒子滤波（Particle Filter, PF）Isard M, Blake A. Condensation - conditional density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29(1): 5-28.相结合来构建跟踪系统。

虽然SIFT是一个较为鲁棒的特征描述子，但其特征维度高，提取实时性差，且存在对边缘光滑的目标无法准确提取特征点、匹配成功特征点数目较少、特征点分布欠均匀等问题。针对这些问题，一些研究者对SIFT特征进行了改进：PCA-SIFT王鹤, 谢刚. 基于PCA-SIFT特征的目标识别算法[J]. 电视技术, 2013, 37(15): 30-32.将SIFT提取算法中的直方图方法换作主元分析法，在图像旋转和光照变化中有较好的性能；SURF（Speeded Up Robust Features）He W, Yamashita T, Lu H, et al. SURF tracking[C]. International Conference on Computer Vision (ICCV), 2009. 1586-1592.包加桐, 宋爱国, 郭晏, et al. 基于SURF特征跟踪的动态手势识别算法[J]. 机器人, 2011, 33(4): 482-489.是SIFT加速版，使用积分图和Hessian矩阵对其进行加速，不仅与SIFT的性能几乎一致，而且还拥有更快的计算速度。

（3）纹理特征

纹理特征度量了图像中物体表面区域的灰度空间分布，描述了方向、粒度、密度、粗糙度、均匀性、规则性等物理属性。纹理特征主要包括Gabor滤波纹理特征Tian Y L, Kanade T, Cohn J F. Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity[C]. IEEE International Conference on Automatic Face and Gesture Recognition, 2002. 229-234.、灰度共生矩阵Haralick R M, Shanmugam K, Dinstein I H. Textural features for image classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1973, 15(6): 610-621.、LBP纹理特征Ojala T, Pietikainen M P, Maenpaa T M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.宋克臣, 颜云辉, 陈文辉, et al. 局部二值模式方法研究与展望[J]. 自动化学报, 2013, 39(6): 730-744.等。在视觉跟踪中应用最广泛的纹理特征就是LBP纹理特征。

LBP纹理特征具有灰度和旋转不变性，并且计算简单，能够有效地描述目标表面变化情况，且对图像噪声有一定的鲁棒性，因此常作为跟踪算法中的一种视觉特征吴刚, 唐振民, 程勇, et al. 灰度共生矩阵纹理特征的运动目标跟踪方法[J]. 南京理工大学学报(自然科学版), 2010, 04): 459-463.李巍, 赵英凯, 钱厚亮. 一种基于纹理和颜色的目标跟踪方法[J]. 计算机仿真, 2011, 28(01): 273-276.Chuan-Xu W, Zuo-Yong L. A new face tracking algorithm based on local binary pattern and skin color information[C]. 2008. 657-660.。文献宁纪锋, 吴成柯. 一种基于纹

理模型的 Mean Shift 目标跟踪算法[J]. 模式识别与人工智能, 2007, 20(5): 612-618.在 Mean Shift 跟踪算法中，使用 LBP 纹理特征作为目标表达方法，利用少量的关键点准确描述目标，计算复杂度较低。在复杂的条件下，该方法比基于颜色的表示法在目标表达的准确性和跟踪性能上均有明显提高。针对红外成像，文献王永忠, 赵春晖, 梁彦, et al. 一种基于纹理特征的红外成像目标跟踪方法[J]. 光子学报, 2007, 36(11): 2163-2167.提出了一种基于 LBP 纹理特征的红外目标跟踪算法，较传统基于灰度的跟踪算法更为鲁棒。

LBP 纹理特征的缺点是容易受光照变化不均的影响。为了解决传统 LBP 特征的缺陷，学者们提出了一些改进的 LBP 纹理特征：局部三值模式（LTP）Tan X Y, Triggs B. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions[J]. IEEE Transactions on Image Processing, 2010, 19(6): 1635-1650.是一种 LBP 的改进特征，采用三值的编码方式，能够有效地解决光照、噪声等问题；局部相位量化（Local Phase Quantization, LPQ）Ojansivu V, Heikkila J. Blur insensitive texture classification using local phase quantization[C]. Lecture notes in computer science, 2008, 236-243.则首先对中心点的四个方向进行频域变换，将频域变换得到的的实部和虚部作为编码值，不仅具有传统 LBP 的特点，而且还具有模糊不变性。

（4）其他局部特征

Haar 特征 Klesk P, Godziuk A, Kapruziak M, et al. Fast analysis of C-Scans from fround penetrating radar via 3-D Haar-Like features with application to landmine detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(7): 3996-4009.反映了图像的灰度变化情况，通过使用积分图进行计算大大地提高了图像特征值计算的效率，但是 Harr 特征对边缘、线段比较敏感。显著性特征（Saliency）Su Y Y, Zhao Q J, Zhao L J, et al. Abrupt motion tracking using a visual saliency embedded particle filter[J]. Pattern Recognition, 2014, 47(5): 1826-1834.是基于显著性检测获取的显著性信息，模仿人类感知机制来获取图像中感兴趣的区域，具有特异性和鲁棒性，但是对噪声和剧烈的光照变化非常敏感。另外，基于分割的特征 Ren X, Malik J. Tracking as repeated figure/ground segmentation[C]. Computer Vision and Pattern Recognition, 2007, 1-8.、Harris 角点特征 Derpanis K G. The Harris Corner Detector[J]. Symposium Svenska Sllskapet Fr Bildanalys, 2004,等也是视觉跟踪领域常用的局部特征。表 3 对上述各局部特征表达进行了归纳。

表 3 局部特征表达分类

Table 3 Classification of Local Feature Expressions

局部特征	参考文献	特点
HOG 特征	[40-47]	对图像的几何和光学形变较鲁棒；计算比较复杂、特征维度高、实时性较差
SIFT 特征	[34][48-54]	具有光照、尺度、仿射等不变性；维度高，实时性差
纹理特征	[46-47][55-63]	具有灰度和旋转不变性，计算简单；容易受光照变化不均的影响

Haar 特征	[64]	计算效率高；对边缘、线段比较敏感
显著性特征	[37]	具有特异性和鲁棒性；对噪声和剧烈的光照变化非常敏感
基于分割的特征	[65]	能够有效处理非刚体目标的外观、尺度、外形等变化
Harris 角点特征	[66]	计算简单，对图像旋转、灰度变化、噪声影响和视点变换不敏感

2.3 全局和局部视觉特征融合

综合 2.1 节和 2.2 节的分析，视觉跟踪中的各种全局视觉特征和局部视觉特征均有各自的优点和缺点。通常，全局视觉特征表示方法较为简单，计算方便快速，但是容易受到由光照变化、形变、旋转和局部遮挡等造成的外观变化影响。相对地，局部视觉特征能够得到目标外观的局部位置、形状、几何结构，因此对于此类外观变化比较鲁棒。但是，由于局部视觉特征往往需要进行关键点检测，而关键点容易受到噪声和背景的干扰，所以局部视觉特征对噪声和背景干扰不太鲁棒。另外，局部视觉特征需要提取图像大量的局部信息，因此其计算量比全局视觉特征大。

视觉跟踪的具体应用环境决定了特定的目标特征表达策略，根据不同的需求可以对不同的全局视觉特征和局部视觉特征进行有效选择与融合，以实现优势特征间的优势互补，从而获得鲁棒的多线索目标特征表达。例如，文献 Zhang B, Tian W, Jin Z. Efficient hybrid appearance model for object tracking with occlusion handling[J]. Machine Vision, Pattern Recognition, 2007, 46(8): 202-213.将目标的颜色特征和纹理特征进行融合，作为跟踪目标的特征描述；文献 Han Z, Ye Q, Jiao J. Combined feature evaluation for adaptive visual object tracking[J]. Computer Vision and Image Understanding, 2011, 115(1): 69-80.将颜色直方图和方向梯度直方图进行融合作为跟踪目标的特征描述；文献 Sun L, Liu G. Visual object tracking based on combination of local description and global representation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(4): 408-420.将全局视觉特征与局部视觉特征融合，建立了混合的目标特征表达。而对于特征融合的方式也有很多，其中 Y Li 等人通过双三次线性插值的方式将 HOG、CN 和灰度特征采样为同一维度，并结合 Li Y, Zhu J. A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration[J]. 2014, 8926:254-265.；G ZHU 等人则是先对原图提取 CN 和 Lab 特征，然后在此基础上再提取 HOG 特征 Zhu G, Wang J, Wu Y, et al. MC-HOG correlation tracking with saliency proposal[C]// Thirtieth AAAI Conference on Artificial Intelligence. AAAI Press, 2016:3690-3696.。

3 数学特征

视觉特征是一种类似于人眼看到的直观特征表达，而与视觉特征不同，数学特征是将图像信息看成一种抽象的数学信号，其目的是构建目标的内在数学联系。常用的数学特征可分为高斯混合模型、子空间模型和 WSL（Wandering、Stable、Lost）混合模型。表 4 对数学特征表达分类进行了归纳。

表 4 数学特征表达分类

Table 4 Classification of Mathematical Feature Expressions

数学特征	参考文献	特点
高斯混合模型	[72-73]	能得到属于每个分类的概率；实际应用中很难恰当选取高斯分量数
子空间模型	[74-79]	计算简单、有效；可能存在高阶冗余信息
WSL 混合模型	[39][80]	可以处理光照变化和外观变形；有时会将变化缓慢的背景估计成目标

3.1 高斯混合模型

高斯混合模型的构建思想是使用一组高斯分布来近似描述目标外观的概率密度函数。高斯混合模型的优势在于，它所得到的结果是样本点属于每个分类的概率，而不是确定的分类标记。

在视觉跟踪中，Han 等 Han B, Davis L. On-Line density-based appearance modeling for object tracking[C]. IEEE International Conference on Computer Vision, 2005. 1492-1499.提出一种使用混合高斯密度函数的目标外观模型，此模型能够应用于实时跟踪场景。Wang 等 Hanzi W, David S, Konrad S, et al. Adaptive object tracking based on an effective appearance filter[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(9): 1661-1667.为了获取被跟踪目标的时空描述，提出了一种空间-颜色混合高斯外观模型（Spatial-color Mixture of Gaussian, SMOG），SMOG 外观模型可以同时对空间信息和颜色信息进行编码，能够在粒子滤波算法框架内实现视觉跟踪。传统基于颜色直方图的相似性度量方法只考虑颜色信息，SMOG 外观模型通过将颜色和布局信息同时融入相似性度量，增强了模型的判别能力。同时，该作者还将边缘点的空间分布、梯度强度也融入到 SMOG 外观模型，进一步增强了 SMOG 外观模型的鲁棒性和稳定性。

关于高斯混合模型的构建，实际应用中往往很难恰当地选取高斯分量数。因此，很多研究者都采用经验设定值或者采用启发式方法来确定，但这样会使得跟踪算法的可扩展性变差。另外，高斯混合模型每一步迭代的计算量比较大，而且有可能陷入局部极值，具体地和初始值的选取密切相关。

3.2 子空间模型

子空间模型是一种简化复杂问题，揭示问题主要矛盾的数学模型。其本质是将原始高维样本投影到一个更有利于分类的低维特征子空间，具有计算简单、有效等特性，在视觉跟踪领域应用广泛。

MichaelBlack M J, Jepson A D. Eigen tracking: robust matching and tracking of articulated objects using a view-based representation[J]. 1996, 26(1): 63-84.是最早使用子空间模型进行视觉跟踪的研究者之一，由于采用离线训练的特征向量来描述目标，因此该算法无法适应目标外观的变化。为了解决这个问题，David RossRoss D, Lim J, Yang M. Adaptive probabilistic visual tracking with incremental subspace update[C]. European Conference on Computer Vision,

2004. 470-482.在此基础上引入了 PCA 子空间，并提出了一种能够对目标特征向量进行在线更新的增量式 PCA 算法，该算法对目标外观的变化具有较好的适应能力。之后，针对 PCA 子空间的改进，相关学者又提出了基于 2D-PCAWang T, Gu I Y H, Shi P. Object tracking using incremental 2D-PCA learning and ML estimation[C]. IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. 933-936. 和鲁棒 PCALeonardis A, Bischof H. Robust recognition using eigenimages[J]. Computer Vision and Image Understanding, 2000, 78(1): 99-118.的子空间跟踪算法。而为了提高了对复杂背景中目标的描述能力，近年来基于张量 Hu W, Li X, Zhang X, et al. Incremental tensor subspace learning and Its applications to foreground segmentation and tracking[J]. International Journal of Computer Vision, 2011, 91(3): 303-327.和黎曼子空间 Li X, Hu W, Zhang Z, et al. Visual tracking via incremental log-euclidean riemannian subspace learning[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008. 1-8.的视觉跟踪算法研究则采用了更加复杂的子空间理论及其更新机制。

由于 PCA 子空间模型在进行数据处理的过程中往往只考虑数据的二阶统计特性，未涉及到其高阶统计信息，因此 PCA 变换后的数据间仍可能存在高阶冗余信息。

3.3 WSL 混合模型

WSL（Wandering、Stable、Lost）混合模型是由 Jepson 等人 Jepson A D, Fleet D J, El-Maraghi T F. Robust online appearance models for visual tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1296-1311.提出的，该模型的优势在于可以处理光照变化和外观变形。WSL 混合模型包含三个分量：Wandering 帧间变分分量（瞬态分量）、Stable 目标稳态分量和 Lost 噪声影响分量。每个分量都采用高斯模型建模，并在跟踪过程中基于在线期望最大算法更新模型参数。Jepson 等人 Jepson A D, Fleet D J, El-Maraghi T F. Robust online appearance models for visual tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1296-1311.是采用滤波的方法来得到目标外观的 WSL 混合模型，而 Zhou 等人 Zhou S, Chellappa R, Moghaddam B. Visual tracking and recognition using appearance-adaptive models in particle filters[J]. IEEE Transactions on Image Processing, 2004, 13(11): 1491-1506.则直接以每个像素的强度作为目标外观的 WSL 混合模型。

在 WSL 混合模型中，主要是通过目标的稳态分量来进行目标的仿射运动估计。因此，WSL 混合模型的主要问题在于，稳态分量有时会将变化缓慢的背景也估计成目标。

4 语义特征

对于一般的图像、视频而言，传统的视觉特征只能表达目标的颜色、形状、纹理等低层次结构特征，因此当目标发生了内在或外在结构变化时，这些视觉特征对目标进行有效描述的能力较差；而数学特征将目标表达为某种过于抽象的数学模型，当目标因为某种外在原因而导致其数学分布发生变化时，数学特征则不能及时适应此变化情况。相比之下，卷积神经网络（Convolutional Neural Network，CNN）作为一种语义特征表达方法，通过

模拟人类神经元的工作过程，让计算机像人类一样慢慢地认识世界，了解物体的内在语义信息。比如一张含有小狗的图像，计算机能够知道的只是图像中一连串的二进制码，而人类却能够通过图像获取“小狗”这种语义信息。语义信息是一个物体所具有的内在属性，当给小狗图像赋予“小狗”这种语义标签后，即使小狗的形状、大小、颜色发生了巨大变化，但是 CNN 仍旧能够根据语义信息来判定这是一只小狗。正是由于 CNN 具有这种强大的提取语义信息的能力，使其在计算机视觉的众多领域均取得了巨大的成功，如：图像识别分类 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science - Research and Development, 2014,He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, 770-778.Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]. International Conference on Neural Information Processing Systems, 2012. 1097-1105.、目标检测 Ouyang W, Wang X, Zeng X, et al. Deepid-net: Deformable deep convolutional neural networks for object detection[C]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015. 2403-2412.Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Computer Vision and Pattern Recognition, 2014, 580-587.、语义分割 Farabet C, Couprie C, Najman L, et al. Learning hierarchical features for scene labeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1915-1929.Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. 2016, 10(79): 1337-1342.以及其他领域 Taigman Y, Yang M, Ranzato M, et al. Deepface: Closing the gap to human-level performance in face verification[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014. 1701-1708.Toshev A, Szegedy C. Deeppose: Human pose estimation via deep neural networks[C]. Computer Vision & Pattern Recognition, 2014. 1653-1660.。

而在视觉跟踪领域，一些基于 CNN 语义特征的跟踪算法也取得了不错的跟踪效果 Wang N, Yeung D. Learning a deep compact image representation for visual tracking[C]. Advances in Neural Information Processing Systems, 2013. 809-817.Hong S, You T, Kwak S, et al. Online tracking by learning discriminative saliency map with convolutional neural network[J]. Computer Science - Research and Development, 2015, 597-606.Wang N, Li S, Gupta A, et al. Transferring rich feature hierarchies for robust visual tracking[J]. Computer Science - Research and Development, 2015,，它们通常采用 CNN 最后一层全连接层的输出特征作为表达跟踪目标的语义特征，这种特征具有良好的语义分辨能力，对于跟踪问题中一些常见的挑战性因素（例如形变、遮挡等）均具有较高的鲁棒性。但是由于 CNN 最后一层所输出的语义信息丢失了过多的空间位置信息，因此它对于待跟踪目标的定位处理往往不够精确。马超 Ma C, Huang J, Yang X, et al. Hierarchical convolutional features for visual tracking[C]. IEEE International Conference on Computer Vision, 2015. 3074-3082.和王利军 Wang L, Ouyang W, Wang X, et al.

Visual tracking with fully convolutional networks[C]. IEEE International Conference on Computer Vision, 2015. 3119-3127.通过研究发现，在 CNN 的浅层（前几层），目标的空间、形状、位置等信息有很好的保留，但是语义信息的描述力度不足。而当 CNN 深度加深时，则可以获得目标更为鲁棒的语义信息，但是又会造成空间结构信息的缺失。因此一些学者 Ma C, Huang J, Yang X, et al. Hierarchical convolutional features for visual tracking[C]. IEEE International Conference on Computer Vision, 2015. 3074-3082.Qi Y, Zhang S, Qin L, et al. Hedged Deep Tracking[C]. IEEE International Conference on Computer Vision and Pattern Recognition, 2016. 4303-4311.Danelljan M, Robinson A, Khan F S, et al. Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking[C]. European Conference on Computer Vision, 2016, 472-488.提出在跟踪过程中同时使用 CNN 提取的浅层空间信息特征和深层语义特征，这样所形成的综合特征不仅可以精确地跟踪定位目标，而且对形变、光照等挑战性因素也具有很高的鲁棒性。

为了提取关于跟踪目标的 CNN 语义特征，就需要非常庞大的样本对 CNN 模型进行训练，而在一般的视觉跟踪任务中，能够获取的训练样本远远达不到 CNN 模型的训练需求。因此一些学者 Oquab M, Bottou L, Laptev I, et al. Learning and transferring mid-level image representations using convolutional neural networks[C]. Computer Vision & Pattern Recognition, 2014. 1717-1724.Tompson J, Goroshin R, Jain A, et al. Efficient object localization using convolutional networks[C]. Computer Vision & Pattern Recognition, 2015. 648-656.提出了使用一个已经在大规模数据库（例如 ImageNetRussakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.）中训练好的预训练 CNN 模型，再结合当前跟踪任务的训练样本，对预训练 CNN 模型进行微调（Fine-Tune）。由于 CNN 语义特征具有很好的泛化性能，因此这种 Fine-Tune 的方法在不同的跟踪任务上均能取得显著效果。王利军 Wang L, Ouyang W, Wang X, et al. Stct: Sequentially training convolutional networks for visual tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016. 1373-1381.将 CNN 和集成学习（Ensemble Learning）联系起来，提出了序贯训练的 CNN 模型，其采用了已经在 VGGnet-D 数据库 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science - Research and Development, 2014,中训练好的 CNN 模型作为预训练 CNN 模型，并创新性地将卷积层的每个通道看成是一个基学习器并逐次训练，最终对所有基学习器训练完毕后便可获得集成学习器。这种方法在视觉跟踪中取得了优异的跟踪效果，但是由于其需要在线地更新 CNN 参数，故其整体实现速度较慢。Martin Danelljan 通过直接将 VGG 网络输出作为相关滤波跟踪算法输入特征实现了 state-of-art 的效果 Danelljan M, Bhat G, Khan F S, et al. ECO: Efficient Convolution Operators for Tracking[J]. 2016.。为了综合深度特征的精准表达和相关滤波类跟踪算法的速度，J Valmadre 等人将相关滤波层嵌入 CNN 网络

进行离线训练，效果有一定的提升，并且达到了实时 Valmadre J, Bertinetto L, Henriques J F, et al. End-to-end representation learning for Correlation Filter based tracking[J]. 2017.。

5 结论

本文将视觉跟踪中常用的目标特征表达方法分为了三类，分别为视觉特征、数学特征和语义特征，并对它们进行了详细地分析和比较。具体地，视觉特征是一种类似于人眼看到的直观特征表达，着重于描述目标外观的时空信息；数学特征是将图像信息看成一种抽象的数学信号，着重于构建目标的内在数学联系；语义特征则是表达物体内在的语义属性是随着卷积神经网络（CNN）在计算机视觉上的广泛应用而提炼、泛化出的一种新型特征。相对于视觉特征和数学特征，语义特征能够更加有效地体现物体内在类别信息，对形变、遮挡等均有较高的鲁棒性。在实际的视觉跟踪任务中，目标特征表达直接关系着最终的跟踪效果，选择合适的目标特征表达方法，将大大提升结果的准确性。

尽管目前关于目标特征表达方法的研究已取得了众多的研究成果，但是在复杂性与鲁棒性、准确性与鲁棒性方面仍然存在如下问题亟需解决：

1）处理目标特征表达在复杂性与鲁棒性之间的关系：研究一种简单而鲁棒的目标特征表达方法是视觉跟踪中的重要问题。一般地，简单的目标特征表达（如灰度特征、CN 特征、颜色直方图、梯度特征、HOG 特征等）其计算量较低，但是跟踪鲁棒性较差。而相对鲁棒的目标特征表达（以 CNN 和 RNN 网络为主的深度特征）虽具有较强的跟踪鲁棒性，但往往计算量又较大。因此，处理好目标特征表达在复杂性与鲁棒性之间的矛盾关系是设计视觉跟踪算法亟需考虑的一个重要问题。

2）处理目标特征表达在准确性和鲁棒性之间的关系：目前常用的目标特征表达方法很难同时保证视觉跟踪的准确性和鲁棒性。针对一些常规或者特定的跟踪场景，为了提高视觉跟踪的准确性，往往需要对多种目标特征进行选择与融合，或者在现有特征的基础上加入若干限制。但是，这样势必会在提高目标特征准确性（场景适应性）的同时降低其在复杂场景下的跟踪鲁棒性，当跟踪目标发生形变或者遮挡时，跟踪效果会急剧变差。因此，在设计目标特征表达方法时，需要同时考虑视觉跟踪的准确性与鲁棒性。

参考文献：

[1] Xiao L, Meng G, Luo H, et al. Dynamic path planning based on improved boundary value problem for unmanned aerial vehicle[J]. Cluster Computing, 2016, 19(4): 1-10.

[2] Wang X, Li B, Geng Q. Runway Detection and Tracking for Unmanned Aerial Vehicle Based on an Improved Canny Edge Detection Algorithm[C]. International Conference on Intelligent Human-Machine Systems and Cybernetics, 2012, 149-152.

[3] Hadi R A, George L E, Mohammed M J. A Computationally Economic Novel Approach for Real-Time Moving Multi-vehicle Detection and Tracking toward Efficient Traffic Surveillance[J]. Arabian Journal for Science & Engineering, 2016, 1-15.

[4] Hai D, Hua T. Development of a tracking-based system for automated traffic data collection for roundabouts[J]. Journal of Modern Transportation, 2017, 1-12.

[5] Joslin C A E Q. Dynamic gesture recognition[C]. Instrumentation and Measurement Technology Conference, 2005. 1706-1711.
[6] Wang N, Shi J, Yeung D, et al. Understanding and diagnosing visual tracking systems[C]. The IEEE International Conference on Computer Vision (ICCV), 2015. 3101-3109.
[7] Wu Y, Lim J, Yang M. Online object tracking: A benchmark[C]. IEEE Conference on Computer Vision and Pattern Recognition((CVPR)), 2013. 2411-2418.
[8] Li X, Hu W M, Shen C H, et al. A survey of appearance models in visual object tracking[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2013, 4(4): 51-58.
[9] 黄国祥. RGB颜色空间及其应用研究[D]. 中南大学, 2002: 88.
[10] 庞晓敏, 闵子建, 阚江明. 基于HSI和LAB颜色空间的彩色图像分割[J]. 广西大学学报(自然科学版), 2011, 36(6): 976-980.
[11] Bradski G. Real time face and object tracking as a component of a perceptual user interface[C]. Applications of Computer Vision, 1998, 214-219.
[12] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]// Computer Vision and Pattern Recognition. IEEE, 2010:2544-2550.
[13] Danelljan M, Khan F S, Felsberg M, et al. Adaptive Color Attributes for Real-Time Visual Tracking[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014:1090-1097.
[14] Choi J, Chang H J, Yun S, et al. Attentional Correlation Filter Network for Adaptive Visual Tracking[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017.
[15] Kengyew S, Kittler J, Petrou M. Defect detection in random colour textures[J]. Image and Vision Computing, 1996, 14(9): 667-683.
[16] 贾慧星, 章毓晋. 基于梯度方向直方图特征的多核跟踪[J]. 自动化学报, 2009, 10): 1283-1289.
[17] 胡铟, 杨静宇. 基于分块颜色直方图的MeanShift跟踪算法[J]. 系统仿真学报, 2009, 5(10): 2936-2939.
[18] 王晓卫, 王旭东, 贺明. 基于直方图比的背景加权的Mean Shift目标跟踪算法[J]. 强激光与粒子束, 2016, 28(05): 19-23.
[19] 陶立超, 赵宇明. 基于分块颜色直方图和粒子滤波的物体跟踪[J]. 计算机工程与应用, 2012, 48(7): 165-168.
[20] Fukunaga K, Hostetler L D. The estimation of the gradient of a density function, with applications in pattern recognition[J]. IEEE Transactions on Information Theory, 1975, 21(1): 32-40.
[21] Birchfield S T, Rangarajan S. Spatiograms versus histograms for region-based tracking[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR), 2005. 1158-1163.
[22] 汪启伟, 万寿红, 岳丽华. 一种新的空间颜色直方图及其度量方法[J]. 小型微型计算机系统, 2014, 35(06): 1338-1341.
[23] 林晖, 张华君. 基于颜色-空间二维直方图的目标跟踪技术研究[J]. 计算机与数字工程, 2009, 37(12): 126-129.
[24] 裴巧娜. 基于光流法的运动目标检测与跟踪技术[D]. 北方工业大学, 2009: 69.
[25] 王亮. 光流技术及其在运动目标检测和跟踪中的应用研究[D]. 国防科学技术大学, 2007: 78.
[26] Nagel H, Enkelmann W. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(5): 565-593.
[27] Kearney J K, Thompson W B, Boley D L. Optical flow estimation: An error analysis of gradient-based methods with local optimization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(2): 229-244.
[28] Weng J, Ahuja N, Huang T S. Matching two perspective views[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(8): 806-825.
[29] Delpiano J, Jara J, Scheer J, et al. Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy[J]. Machine Vision and Applications, 2012, 23(4SI): 675-689.
[30] 胡庭波, 吴涛, 贺汉根. 基于立体匹配技术的光流场计算方法[J]. 计算机工程与科学, 2006, 28(10): 50-53.
[31] Smith S M. ASSET-2: real-time motion segmentation and shape tracking[C]. Computer

Vision, 1995. Proceedings, Fifth International Conference on, 1995. 80-85.
[32] Shi J, Tomasi C. Good features to track[C]. IEEE Computer Vision and Pattern Recognition(CVPR), 1994. 593-600.
[33] 吴垠, 李良福, 肖樟树, et al. 基于尺度不变特征的光流法目标跟踪技术研究[J]. 计算机工程与应用, 2013, 49(15): 157-161.
[34] Lowe D G. Object recognition from local scale-invariant features[C]. Seventh IEEE International Conference on Computer Visionthe, 1999. 1150-1157.
[35] 陈添丁, 胡鉴, 吴涤. 稀疏光流快速计算的动态目标检测与跟踪[J]. 中国图象图形学报, 2013, 18(12): 1593-1600.
[36] Bowyer K, Kranenburg C. Edge detector evaluation using empirical ROC curves[C]. IEEE Computer Vision and Pattern Recognition(CVPR), 1999. 100-105.
[37] Su Y Y, Zhao Q J, Zhao L J, et al. Abrupt motion tracking using a visual saliency embedded particle filter[J]. Pattern Recognition, 2014, 47(5): 1826-1834.
[38] Sun X, Yao H, Zhang S. A novel supervised level set method for non-rigid object tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011. 3393-3400.
[39] Jepson A D, Fleet D J, El-Maraghi T F. Robust online appearance models for visual tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1296-1311.
[40] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. IEEE Computer Vision and Pattern Recognition(CVPR), 2005. 886-893.
[41] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.
[42] Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking[C]. the British Machine Vision Conference, 2014. 524-533.
[43] 田仙仙, 鲍泓, 徐成. 一种改进HOG特征的行人检测算法[J]. 计算机科学, 2014, 41(09): 320-324.
[44] Wu J, Yang S, Zhang L. Pedestrian detection based on improved HOG feature and robust adaptive boosting algorithm[C]. 2011 4th International Congress on Image and Signal Processing (CISP), 2011. 1535-1539.
[45] Sanin A, Sanderson C, Harandi M T, et al. K-tangent spaces on riemannian manifolds for improved pedestrian detection [C]. IEEE International Conference on Image Processing ICIP, 2012. 473-476.
[46] Ojala T, Pietikainen M P, Maenpaa T M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.
[47] 宋克臣, 颜云辉, 陈文辉, et al. 局部二值模式方法研究与展望[J]. 自动化学报, 2013, 39(6): 730-744.
[48] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
[49] Zhou H, Yuan Y, Shi C. Object tracking using SIFT features and mean shift[J]. Computer Vision and Image Understanding, 2009, 113(3): 345-352.
[50] Fazli S, Pour H M, Bouzari H. Particle Filter based Object Tracking with Sift and Color Feature[C]. International Conference on Machine Vision (ICMV), 2009. 89-93.
[51] Isard M, Blake A. Condensation - conditional density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29(1): 5-28.
[52] 王鹤, 谢刚. 基于PCA-SIFT特征的目标识别算法[J]. 电视技术, 2013, 37(15): 30-32.
[53] He W, Yamashita T, Lu H, et al. SURF tracking[C]. International Conference on Computer Vision (ICCV), 2009. 1586-1592.
[54] 包加桐, 宋爱国, 郭晏, et al. 基于SURF特征跟踪的动态手势识别算法[J]. 机器人, 2011, 33(4): 482-489.
[55] Tian Y L, Kanade T, Cohn J F. Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity[C]. IEEE International Conference on Automatic Face and Gesture Recognition, 2002. 229-234.
[56] Haralick R M, Shanmugam K, Dinstein I H. Textural features for image classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1973, 15(6): 610-621.

[57] 吴刚，唐振民，程勇，et al. 灰度共生矩阵纹理特征的运动目标跟踪方法[J]. 南京理工大学学报(自然科学版)，2010，04)：459-463.
[58] 李巍，赵英凯，钱厚亮. 一种基于纹理和颜色的目标跟踪方法[J]. 计算机仿真，2011，28(01)：273-276.
[59] Chuan-Xu W, Zuo-Yong L. A new face tracking algorithm based on local binary pattern and skin color information[C]. 2008. 657-660.
[60] 宁纪锋，吴成柯. 一种基于纹理模型的 Mean Shift 目标跟踪算法[J]. 模式识别与人工智能，2007，20(5)：612-618.
[61] 王永忠，赵春晖，梁彦，et al. 一种基于纹理特征的红外成像目标跟踪方法[J]. 光子学报，2007，36(11)：2163-2167.
[62] Tan X Y, Triggs B. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions[J]. IEEE Transactions on Image Processing, 2010, 19(6): 1635-1650.
[63] Ojansivu V, Heikkila J. Blur insensitive texture classification using local phase quantization[C]. Lecture notes in computer science, 2008, 236-243.
[64] Klesk P, Godziuk A, Kapruziak M, et al. Fast analysis of C-Scans from fround penetrating radar via 3-D Haar-Like features with application to landmine detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(7): 3996-4009.
[65] Ren X, Malik J. Tracking as repeated figure/ground segmentation[C]. Computer Vision and Pattern Recognition, 2007, 1-8.
[66] Derpanis K G. The Harris Corner Detector[J]. Symposium Svenska Sllskapet Fr Bildanalys, 2004,
[67] Zhang B, Tian W, Jin Z. Efficient hybrid appearance model for object tracking with occlusion handling[J]. Machine Vision, Pattern Recognition, 2007, 46(8): 202-213.
[68] Han Z, Ye Q, Jiao J. Combined feature evaluation for adaptive visual object tracking[J]. Computer Vision and Image Understanding, 2011, 115(1): 69-80.
[69] Sun L, Liu G. Visual object tracking based on combination of local description and global representation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(4): 408-420.
[70] Li Y, Zhu J. A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration[J]. 2014, 8926:254-265.
[71] Zhu G, Wang J, Wu Y, et al. MC-HOG correlation tracking with saliency proposal[C]// Thirtieth AAAI Conference on Artificial Intelligence. AAAI Press, 2016:3690-3696.
[72] Han B, Davis L. On-Line density-based appearance modeling for object tracking[C]. IEEE International Conference on Computer Vision, 2005. 1492-1499.
[73] Hanzi W, David S, Konrad S, et al. Adaptive object tracking based on an effective appearance filter[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(9): 1661-1667.
[74] Black M J, Jepson A D. Eigen tracking: robust matching and tracking of articulated objects using a view-based representation[J]. 1996, 26(1): 63-84.
[75] Ross D, Lim J, Yang M. Adaptive probabilistic visual tracking with incremental subspace update[C]. European Conference on Computer Vision, 2004. 470-482.
[76] Wang T, Gu I Y H, Shi P. Object tracking using incremental 2D-PCA learning and ML estimation[C]. IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. 933-936.
[77] Leonardis A, Bischof H. Robust recognition using eigenimages[J]. Computer Vision and Image Understanding, 2000, 78(1): 99-118.
[78] Hu W, Li X, Zhang X, et al. Incremental tensor subspace learning and Its applications to foreground segmentation and tracking[J]. International Journal of Computer Vision, 2011, 91(3): 303-327.
[79] Li X, Hu W, Zhang Z, et al. Visual tracking via incremental log-euclidean riemannian subspace learning[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008. 1-8.
[80] Zhou S, Chellappa R, Moghaddam B. Visual tracking and recognition using appearance-adaptive models in particle filters[J]. IEEE Transactions on Image Processing, 2004, 13(11): 1491-1506.
[81] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image

recognition[J]. Computer Science - Research and Development, 2014,
[82] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, 770-778.
[83] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]. International Conference on Neural Information Processing Systems, 2012. 1097-1105.
[84] Ouyang W, Wang X, Zeng X, et al. Deepid-net: Deformable deep convolutional neural networks for object detection[C]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015. 2403-2412.
[85] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Computer Vision and Pattern Recognition, 2014, 580-587.
[86] Farabet C, Couprie C, Najman L, et al. Learning hierarchical features for scene labeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1915-1929.
[87] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. 2016, 10(79): 1337-1342.
[88] Taigman Y, Yang M, Ranzato M, et al. Deepface: Closing the gap to human-level performance in face verification[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014. 1701-1708.
[89] Toshev A, Szegedy C. Deeppose: Human pose estimation via deep neural networks[C]. Computer Vision & Pattern Recognition, 2014. 1653-1660.
[90] Wang N, Yeung D. Learning a deep compact image representation for visual tracking[C]. Advances in Neural Information Processing Systems, 2013. 809-817.
[91] Hong S, You T, Kwak S, et al. Online tracking by learning discriminative saliency map with convolutional neural network[J]. Computer Science - Research and Development, 2015, 597-606.
[92] Wang N, Li S, Gupta A, et al. Transferring rich feature hierarchies for robust visual tracking[J]. Computer Science - Research and Development, 2015,
[93] Ma C, Huang J, Yang X, et al. Hierarchical convolutional features for visual tracking[C]. IEEE International Conference on Computer Vision, 2015. 3074-3082.
[94] Wang L, Ouyang W, Wang X, et al. Visual tracking with fully convolutional networks[C]. IEEE International Conference on Computer Vision, 2015. 3119-3127.
[95] Qi Y, Zhang S, Qin L, et al. Hedged Deep Tracking[C]. IEEE International Conference on Computer Vision and Pattern Recognition, 2016. 4303-4311.
[96] Danelljan M, Robinson A, Khan F S, et al. Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking[C]. European Conference on Computer Vision, 2016, 472-488.
[97] Oquab M, Bottou L, Laptev I, et al. Learning and transferring mid-level image representations using convolutional neural networks[C]. Computer Vision & Pattern Recognition, 2014. 1717-1724.
[98] Tompson J, Goroshin R, Jain A, et al. Efficient object localization using convolutional networks[C]. Computer Vision & Pattern Recognition, 2015. 648-656.
[99] Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
[100] Wang L, Ouyang W, Wang X, et al. Stct: Sequentially training convolutional networks for visual tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016. 1373-1381.
[101] Danelljan M, Bhat G, Khan F S, et al. ECO: Efficient Convolution Operators for Tracking[J]. 2016.
[102] Valmadre J, Bertinetto L, Henriques J F, et al. End-to-end representation learning for Correlation Filter based tracking[J]. 2017.